Automated Ptosis Measurements From Facial Photographs

Zachary M. Bodnar, MD; Michael Neimkin, MD; John B. Holds, MD

The standard clinical evaluation of ptosis includes manual measurement of eyelid positions to quantify the degree of ptosis and its effect on a patient’s vision. This value is most often expressed as the margin reflex distance 1 (MRD1), defined by the vertical distance between the upper eyelid margin and the corneal light reflex, which is the specular reflection at the corneal apex from a light source that is aligned with the visual axis (eg, a penlight). Similarly, the margin reflex distance 2 (MRD2) is the vertical distance between the corneal light reflex and the lower eyelid margin. Another way that ptosis is commonly quantified is to plot the area of visual field deprivation that results from eyelid malposition by performing Goldmann perimetry, once with the eyelid in its natural position and again with the eyelid taped up to simulate the results of surgical correction. Problems that limit the usefulness of these methods include operator dependence, subjectivity, patient movement, and cognitive inability to participate in testing (eg, in children or cognitively impaired adults). The capability to objectively extract measurements of the MRD1 and MRD2 from a single photograph of a patient, obtained in the clinic with an inexpensive digital camera, could mitigate many of these difficulties.

A software algorithm developed by one of us (Z.M.B.) can identify the corneal light reflex and eyelid margins in frontal photographs of human faces and automatically calculate the MRD1 and MRD2. The objectives of this study were to introduce our method for computing automated ptosis measurements from digital photographs and to compare it with standard manual clinical measurements.

**Software Algorithm**

Our software algorithm accepts as input a single digital photograph in any one of several standard file formats (JPEG, TIFF, etc) and an optional set of calibration matrices as follows:

\[
C = \begin{bmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix}
\]
and $K = [k_3, k_2, k_3, k_4, k_5]$, where $f_x$ and $f_y$ are focal lengths, $c_x$ and $c_y$ is the image center in pixels, and $[k_i]$ are radial distortion coefficients specific to the camera used to acquire the image. These matrices may be used to remove any significant distortion introduced by the camera lens using a method previously described by Zhang. Our software algorithm can be used to estimate the matrices $C$ and $K$ for a particular camera. First, a set of at least 10 checkerboard-pattern images in different orientations is acquired by the camera undergoing calibration. These images are analyzed by the software algorithm to estimate $C$ and $K$ using another well-known computer vision algorithm. 

The camera calibration method is required only once per lens and camera combination of a particular design and manufacture and may not be necessary for high-quality lenses that introduce little radial distortion. EFigure 2 in the Supplement shows the results of camera calibration and undistortion on a lens with a large amount of radial distortion.

**Image Analysis**

The undistorted image is segmented using standard face detection algorithms to first identify a region corresponding to the face. A search is performed within this region to identify the eyes. Facial feature edges are then detected within the regions corresponding to the eyes using the standard edge detection algorithm developed by Canny. The corneal light reflex of each eye is determined by the specular reflection of the camera flash, and is identified by the software algorithm as the area enclosed by the edge nearest to the highest-intensity pixel of the corresponding eye region. In the vertical dimension, the distances from the centroid of the corneal light reflex to the detected feature edges are reported.

To enable the software algorithm to determine the scale of the image plane in millimeters per pixel, a circular marker of known radius (in millimeters) is placed on the individual’s forehead before image capture. The marker is automatically detected, and its radius (in image pixels) is determined using another well-known computer vision algorithm. Figure 1 shows a sample input photograph and the software algorithm’s output, consisting of a re-rendering of the original photograph with the detected edges and calculated measurements overlaid. Processing one such photograph takes only a few seconds.

**Methods**

This observational study was performed at a single-surgeon (J.B.H.) oculoplastic private practice. The study dates were July 30, 2014, to September 12, 2014. The dates of our analysis were October 22, 2014, to June 18, 2015. Clinical photographs were obtained at 4288 × 2848 (12.3 megapixel) resolution using a commercially available camera system (Figure 2). The software algorithm was implemented using the Java programming language and an open-source computer vision library (OpenCV, version 2.4.6; Itseez). Institutional review board approval of the study was obtained from Saint Louis University. Written informed consent was obtained from study participants in accord with the Declaration of Helsinki, and the study was performed with Health Insurance Portability and Accountability Act of 1996 compliance. 

**Results**

In total, 55 eyes of 28 volunteers 24 to 80 years old (mean age, 57 years) were included in the study (one eye with an absent corneal light reflex because of a negative MRD1 was excluded). Five participants (18%) were male, and 23 participants (82%) were female. Twenty-six participants (93%) were of white race/ethnicity, and 2 participants (7%) were African American. The manual MRD1 measurements ranged from 0 to 6.00 mm (mean [SD], 2.87 [1.56] mm). The automated MRD1 measurements ranged from 0.23 to 5.64 mm (mean [SD], 2.91 [1.48] mm). The manual MRD2 measurements ranged from 4.00 to 9.00 mm (mean [SD], 5.68 [1.19] mm). The automated MRD2 measurements ranged from 4.34 to 9.46 mm (mean [SD], 5.81 [1.19] mm).

Manual measurements of the MRD1 and MRD2 in each individual were obtained by us using a penlight to produce the corneal light reflex and a ruler to judge the distance. A single camera system (camera, lens, and flash model) was used to obtain frontal facial photographs of all study participants at the time of manual measurements at a fixed focal length of 1 m, with the camera positioned directly in front of the patient at eye level at a distance of 1 m as determined by the focal plane of the camera. At this distance, small translational and rotational variation in camera position should have minimal effect on the corneal light reflex and eyelid measurements. These variables were chosen to simulate image acquisition in a realistic clinical setting.

The camera was calibrated using 16 images. A marker with a diameter of 19.05 mm was placed on the forehead of each individual to provide a reference of scale. The images were processed using our software algorithm, and the MRD1 and MRD2 measurements for each photograph were recorded. Because the large distance from the optical center of the lens to the camera flash (108.00 mm) causes some displacement in the corneal light reflex, measurements from photographs with the flash positioned above, below, to the left of, and to the right of the camera lens were averaged for comparison with manual measurements and with each other to assess the effect of the flash position. Statistical analyses were performed using a software package (R, version 3.1.2; The R Foundation). Bland-Altman analysis was used to compare agreement between the manual measurements and the automated photometric measurements. $P < .05$ was considered statistically significant. The $r$ values are by Pearson product moment correlation.
Automated MRD1 measurements were normally distributed according to the Shapiro-Wilk test ($P = .05$). They were strongly correlated with manual measurements ($r = 0.97$; 95% CI, $r = 0.95$ to $r = 0.98$; $P < .001$). The bias of automated MRD1 measurements was 0.03 mm (95% CI, −0.06 to 0.12 mm), with 95% confidence limits of −0.66 and 0.71 mm as determined by Bland-Altman analysis. Figure 3 shows a Bland-Altman analysis of the MRD1 results.

Automated MRD2 measurements were normally distributed ($P = .01$). They were strongly correlated with manual measurements ($r = 0.96$; 95% CI, $r = 0.93$ to $r = 0.98$; $P < .001$). The bias of automated MRD2 measurements was 0.13 mm (95% CI, 0.03-0.22 mm), with 95% confidence limits of −0.54 and 0.80 mm. Figure 4 shows a Bland-Altman analysis of the MRD2 results.

Displacement of the corneal light reflex occurred as a result of the distance between the center of the camera lens and the flash. The effect is asymmetric when the flash is oriented horizontally, with the corneal light reflex being more off center than that of the right eye. The numerals in B are in millimeters.

Photometric measurements were obtained with the software algorithm. $\sigma$ indicates SD; $\mu$, arithmetic mean.
compared the differences between measurements made with the flash in the 4 different positions with each other and the averaged measurements and then evaluated these variations for statistical significance using a 2-tailed t test.

When in a horizontal plane, the effect of the flash position was minimal. The mean absolute difference between MRD1 measurements with the flash in the left and right positions was 0.23 mm, which was not statistically significant (P = .69). The mean absolute difference between MRD1 measurements with the flash in the left position and the averaged measurements was −0.31 mm OD and −0.10 mm OS, which was statistically significant for right eyes only (P < .001 and P = .23, respectively). The mean absolute difference between MRD1 measurements with the flash in the right position and the averaged measurements was −0.30 mm OD and −0.20 mm OS, which was statistically significant for right eyes only (P = .002 and P = .23, respectively). The mean absolute difference between MRD2 measurements with the flash in the left position and the averaged measurements was 0.25 mm and was not statistically significant (P = .32). The mean absolute difference between MRD2 measurements with the flash in the left position and the averaged measurements was 0.23 mm OD and 0.25 mm OS, with neither difference being statistically significant (P = .09 and P = .32, respectively).

Vertical flash placement causes significant variation in measurements. Positioning the flash superiorly underestimated the MRD1 by a mean of 0.53 mm (P = .001) and overestimated the MRD2 by a mean of 0.43 (P = .004). The Table summarizes these results.

### Table. Effect of the Flash Position on the Measurement of MRD1 and MRD2

<table>
<thead>
<tr>
<th>MRD1, mm</th>
<th>P Value</th>
<th>MRD2, mm</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δ(Left, right)</td>
<td>0.23</td>
<td>.69</td>
<td>0.25</td>
</tr>
<tr>
<td>Δ(Left, mean)</td>
<td>OD -0.31</td>
<td>&lt;.001</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>OS -0.1</td>
<td>.23</td>
<td>0.2</td>
</tr>
<tr>
<td>Δ(Right, mean)</td>
<td>OD -0.3</td>
<td>.002</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>OS -0.2</td>
<td>.23</td>
<td>-0.01</td>
</tr>
<tr>
<td>Δ(Up, mean)</td>
<td>-0.53</td>
<td>.001</td>
<td>0.42</td>
</tr>
<tr>
<td>Δ(Down, mean)</td>
<td>0.86</td>
<td>&lt;.001</td>
<td>-0.96</td>
</tr>
</tbody>
</table>

Abbreviations: MRD1, margin reflex distance 1; MRD2, margin reflex distance 2.

**Discussion**

Manual measurements of the MRD1 and MRD2 are used in the clinical evaluation of ptosis and the surgical planning of ptosis repair. In addition, these measurements, along with photographic documentation of ptosis, are typically required by insurers to prove medical necessity. However, interobserver variability, reproducibility, patient movement, and poor cooperation with testing present a challenge to current methods of preoperative evaluation. In conjunction with the MRD1 and MRD2, additional considerations in aesthetic eyelid surgery include measurements of the brow fat span and tarsal platform show, as described by Goldberg and Lew.6 Because the relevant facial features that define these metrics can be identified by our method of edge detection analysis, indirect measurements of the brow fat span and tarsal platform show can be obtained from our software algorithm’s output.

Computer-assisted analysis of facial photographs for measurement of the MRD1, MRD2, eyelid contour, and palpebral fissure has been previously described.8-10 However, the methods used rely on significant user and computer interaction after image acquisition and depend on an observer to identify edges and facial features. To our knowledge, ours is the first software algorithm capable of obtaining eyelid measurements with completely automated image processing.

One key issue was displacement of the corneal light reflex by the position of the flash relative to the lens. For data collection in this study, we chose to average readings from all 4 quadrants (superior, left, right, and inferior). In clinical practice, one would most likely choose to place the flash as close to the center of the lens as possible or position it to the left or right of the lens for all photographs. In our data set, choosing only one position altered the measurements by approximately 0.20 mm, which is well
within the margin of error for clinical practice. Therefore, a single photograph with the flash positioned in the horizontal plane would prove adequate for routine clinical measurement. Accuracy could be further improved by minimizing the distance between the flash and the optical elements of the camera.

The 95% confidence limits of our method are within 0.71 mm for MRD1 and 0.80 mm for MRD2, with a bias of 0.13 mm for MRD2 measurements. To be considered equivalent to manual measurements, it would be preferable that the 95% confidence limits are within 0.50 mm. One reason for disagreement may include a difference in the reflex position secondary to displacement of the flash from the optical axis. When there is a high-contrast transition from the pupil to the iris, the pupillary border can be delineated by the standard edge detection algorithm used by our software algorithm, as shown in Figure 3. However, at this stage of development, pupil detection is not reliable. Therefore, we rely on the reflection of the flash as a basis for measurements. This necessity is a practical limitation of our present software algorithm implementation because the use of the pupil centroid in lieu of the corneal light reflex would render the analysis in the Table unnecessary and simplify the measurements. Another limitation is that negative MRD1 and MRD2 values cannot be estimated from a single photograph (when the corneal light reflex or pupil center would be obscured). Manifest strabismus will also cause significant displacement of the corneal light reflex and is an additional limitation, which may be overcome in some cases by using multiple photographs if the patient is capable of alternate fixation.

Some disagreement may also result from placement of the scale marker in the plane of the forehead, which is slightly anterior to the plane of the eyelids. Estimating the scale by using white-to-white measurements of the corneal limbus from the photographs compared with the mean corneal diameter could eliminate the need for the scale marker and might influence the bias or disagreement of the automated measurements. Finally, disagreement between the software algorithm and the manual measurements may simply be due to the limited precision of the manual measurements, which is 0.50 mm.

The participants in our study were recruited and photographed during a single routine clinic visit. Therefore, variability of successive measurements (manual and automated) was not quantified by this study but is an important factor to consider before our method could be used in clinical practice. All individuals were photographed with the head in frontal position. Additional measurements from different positions may help assess the degree to which variability of head position and patient movement degrades the reliability of measurements using our software algorithm. Further work is necessary to render our software algorithm reliable in the routine clinical setting because lighting, facial positioning, and other variables can affect the photographic appearance of ptosis and the results of digital image processing.11 Given these limitations, we do not expect our results to change clinical practice in the near future, and manual measurement of the MRD1 and MRD2 should remain the standard of care. However, we believe that we have demonstrated a proof of concept for a potentially useful new approach to ptosis measurements.

Conclusions

An automated, photography-based system could provide an archival and reproducible means for obtaining the MRD1, MRD2, and other facial morphometric data while mitigating potential sources of error, including movement and observer bias. Automated ptosis measurements produced by our software algorithm compare favorably with manual clinical measurements. Software algorithm-based correction of lens distortion could make this technology available in inexpensive handheld devices, including smartphones.